Chapter 6 Exercise 4

Allen Church

11/6/2019

require(knitr)  
require(haven)  
  
opts\_chunk$set(echo = TRUE)  
options(digits = 3)  
  
load("Ch6\_Exercise4\_Speeding\_tickets.RData")

4a. Difference of means test to assess if fines for men and women are different. Do we have reason to expect endogeneity?

diff\_means\_ols <- lm(dta$Amount ~ dta$Female)  
summary(diff\_means\_ols)

##   
## Call:  
## lm(formula = dta$Amount ~ dta$Female)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -121.7 -49.7 -6.7 30.3 600.3   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 124.665 0.386 323.2 <2e-16 \*\*\*  
## dta$Female -7.939 0.670 -11.8 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 56.1 on 31672 degrees of freedom  
## (36683 observations deleted due to missingness)  
## Multiple R-squared: 0.00442, Adjusted R-squared: 0.00438   
## F-statistic: 140 on 1 and 31672 DF, p-value: <2e-16

The results above indicate a predicted value of $124.66 fine for women. Beta 1 indicates that men (as the treatment group) are associated with a -7.939 decrease in fine amount. A t statistic of 11.8 and p value is approximately 0 indicates this is statistically significant. However, we may suspect endogeneity as the model does not account for race of the driver. Namely, the race of the driver may influence the ticket amount and can be correlated with whether the driver is male or female.

4b. Difference of means test for men and women that controls for age and miles per hour. Do we have reason to expect endogeneity?

diff\_means\_ols2 <- lm(dta$Amount ~ dta$Female + dta$Age + dta$MPHover)  
summary(diff\_means\_ols2)

##   
## Call:  
## lm(formula = dta$Amount ~ dta$Female + dta$Age + dta$MPHover)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -309.39 -20.37 6.84 25.50 225.61   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.8986 0.9716 5.04 4.6e-07 \*\*\*  
## dta$Female -3.5552 0.4737 -7.51 6.3e-14 \*\*\*  
## dta$Age 0.0258 0.0176 1.47 0.14   
## dta$MPHover 6.8769 0.0387 177.67 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 39.6 on 31670 degrees of freedom  
## (36683 observations deleted due to missingness)  
## Multiple R-squared: 0.503, Adjusted R-squared: 0.503   
## F-statistic: 1.07e+04 on 3 and 31670 DF, p-value: <2e-16

After controlling for age and miles per hour, the model suggests that miles per hour is highly statistically significant with a t statistic of 177.67 and a p value of approximately 0. However, we may suspect endogeneity as the model does not account for type of car. Namely, the type of car (i.e. if it is an expensive sports car) can be correlated with the fine amount, whether the driver is male or female, age of the driver, and miles per hour over the speed limit.

4c. Are fines higher for African Americans and Hispanics compared to everyone else? Explain what the coefficients on these variables mean?

diff\_means\_ols3 <- lm(dta$Amount ~ dta$Female + dta$Black + dta$Hispanic)  
summary(diff\_means\_ols3)

##   
## Call:  
## lm(formula = dta$Amount ~ dta$Female + dta$Black + dta$Hispanic)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -121.0 -49.0 -6.3 31.0 594.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 123.963 0.403 307.49 < 2e-16 \*\*\*  
## dta$Female -7.681 0.671 -11.45 < 2e-16 \*\*\*  
## dta$Black 5.620 1.439 3.91 9.4e-05 \*\*\*  
## dta$Hispanic 7.065 1.495 4.73 2.3e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 56.1 on 31670 degrees of freedom  
## (36683 observations deleted due to missingness)  
## Multiple R-squared: 0.00554, Adjusted R-squared: 0.00544   
## F-statistic: 58.8 on 3 and 31670 DF, p-value: <2e-16

The coefficient for African Americans indicates an estimate of 6.25 increase in fine for tickets for African Americans. The t statistic of 4.34 indicates this is statistically significant. The coefficient for Hispanics indicates an estimate of $7.98 increase in fine for tickets for Hispanics The t statistic of 5.34 indicates this is statistically significant. The standard error on these coefficients measure how much Beta One Hat will vary.

4d. The standard error for Female, Black, and Hispanic variables are 0.670, 1.441, and 1.496, respectively. The standard error for Black and Hispanic variables is higher because the variance of these coefficients are higher. Therefore, we have a less precise view for these two variables due to their increased error terms.

4e. Within a single model, assess whether miles over speed limit has differential effect on fines for women, African Americans, and Hispanics

dta$interact\_black <- dta$Black \* dta$MPHover  
dta$interact\_hispanic <- dta$Hispanic \* dta$MPHover  
diff\_means\_ols4 <- lm(dta$Amount ~ dta$Female + dta$MPHover + dta$interact\_black + dta$interact\_hispanic)  
summary(diff\_means\_ols4)

##   
## Call:  
## lm(formula = dta$Amount ~ dta$Female + dta$MPHover + dta$interact\_black +   
## dta$interact\_hispanic)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -308.28 -20.09 6.77 25.60 226.02   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.8885 0.7203 8.17 3.1e-16 \*\*\*  
## dta$Female -3.4914 0.4743 -7.36 1.9e-13 \*\*\*  
## dta$MPHover 6.8617 0.0391 175.54 < 2e-16 \*\*\*  
## dta$interact\_black -0.0585 0.0530 -1.10 0.26970   
## dta$interact\_hispanic 0.1964 0.0559 3.52 0.00044 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 39.6 on 31669 degrees of freedom  
## (36683 observations deleted due to missingness)  
## Multiple R-squared: 0.504, Adjusted R-squared: 0.504   
## F-statistic: 8.03e+03 on 4 and 31669 DF, p-value: <2e-16

The coefficients above suggest MPH over has a positive differential effect for women, African Americans, and Hispanics. The coefficient for women and MPH over suggest that women are associated with a 12.74 dollars increase in ticket amount for each MPH over the speed limit. The coefficients suggest that African American men are associated with a 9.20 dollars increase in ticket amount for each MPH over the speed limit. The coefficients suggest that Hispanic men are associated with a 9.44 dollars increase in ticket amount for each MPH over the speed limit.